1. SEGMENTAL PHONOLOGY

1.1 Writing System

Most of the accurate linguistic work on Mbay and other Sara Languages has been done by linguists at the Centre d'Etudes Linguistiques of the Collège Charles Lwanga (CCL) at Sarh. One major achievement of CCL has been the development of a writing system. This system, based on IPA phonetic symbols, allows for accurate transcription of the phonemes, tones and canonical patterns of these languages. The system of transcription found in this work follows very closely the CCL orthography. Magriña and Fournier (1974), with its careful transcription of Mbay tones and vowel length contrasts, has been especially useful.

1.2 Segment Inventory and Pronunciation

At the phonemic level, there are seven vowels and seventeen consonants in Mbay. In addition, there are four prenasalized stops. A brief description of the pronunciation of these phonemes follows:

1.2.1 Consonants:

The following consonants occur in Mbay:

(1)

	Stops		Prenas	Impl	Nasal	Liq	Glide	Other
Bilabial	b	p	mb	ɓ	m			
Alveolar	d	t	nd	ɗ	n	r,l		s
Pal/post-alv	--	--	nj	--	[ɲ]	--	y	j
Velar	g	k	ŋg	--	[ŋ]		w	
Other							h	

The velar /ŋg/ will be written /ng/ for convenience. The letters /b/, /p/, /j/ and /s/ are pronounced very similar to the way they are pronounced in English. The /d/ and /t/ are pronounced forward on the alveolar ridge (towards the top teeth) than in English, but they are not dental. The velars /k/ and /g/ are similar to their counterparts in English, but both are pronounced with somewhat greater force, especially before schwa /ə/. The /r/ is a brief trill, shorter than Spanish /rr/ but longer than Spanish /r/. For some speakers, /r/ might be described as a lateral flap.[1] The obstruents (b,p,d,t,ɓ,ɗ,k,g,j,s,h,mb,nd,nj,ng) do not occur at the end of a word or at the end of a syllable, while the sonorant and liquid consonants (n,m,ŋ,l,r,y,w) do.

The palatal nasal [ɲ] occurs only in word initial position when the following consonant is nasalized, and hence does not contrast with /y/ which never occurs in this position:

(2)

yàbè	**hippopotamus**	yà̰	[ɲà̰]	**to leave**
yòo	**death**	yō̰o̰	[ɲō̰o̰]	**termite**
yā	**hide**	yá̰a̰	[ɲá̰a̰]	**thing**

Hence, the phoneme /y/ will be used to represent this sound. Both /y/ and /r/ become heavily nasalized when adjacent to a nasalized vowel at the end of a syllable.[2]

Note that I differ in this respect from the system used by CCL, which has chosen instead to posit two additional phonemes, nasalized /y̰/ and /r̰/, and to further claim that vowels adjacent to this phoneme are always nasalized. The disadvantages to their approach are numerous: 1), it fails to explain why /y/ never occurs adjacent to a nasalized vowel; 2) it fails to explain why /o/ and /ɔ/ can contrast when following a nasal (e.g. /mòtè/ 'sick' vs. /mɔ̀tè/ 'penis'), but do not contrast when following /y̰/; 3) two additional phonemes, both unusual within the languages of the world, need to be added to the phonemic inventory; 4) the claim that vowels are always heavily nasalized after nasals is not completely correct. In bisyllabic roots beginning with nasals, nasalization does not occur (e.g. /nàtè/ [nà·tè] 'antelope' vs. /nà̰ā̰-té/ 'together'). More importantly, the approach implies that all vowels are nasalized after nasals, and hence that they are neutralized in this environment. But this is simply not true. Thus, /nàr/ 'money' contrasts with /Nà̰r/, the 'the Nar people, /màr/ 'crocodile' contrasts with /mà̰r/ 'insist', and /màý/ 'inside out' contrasts with /mà̰y/ 'resist [death]'.

1.2.1.1 Prenasalized Stops

There are four prenasalized stops in Mbay (mb,nd,nj,ng). When beginning a word, the nasal in such sequences is not a separate syllable, and contrasts with nasal prefixes added to words beginning with obstruents (e.g. /mbā/ 'propose' vs. /m̄-bā/ 'I pay', /ndà/ 'hit' vs. /ǹ-dà/ 'he show'). The nasal in prenasalized stops is pronounced almost simultaneously with the obstruent, and can seem inaudible to the inexperienced ear.

The /n/ in /nj/ is not palatalized, but the nasal portion of /ng/ is always pronounced [ŋ]. Prenasalized stops behave as obstruents in that they do not occur in final position. In word final position, /ng/ is pronounced [ŋ]:

(3)

wōng	[wōŋ]	**anger**
dāng	[dāŋ]	**misery**
gán̄g	[gáŋ̄]	**winnow**

When /ng/ occurs between vowels, the /g/ is normally pronounced, although some speakers drop it, in which case the vowel which follows becomes heavily nasalized (e.g. /jíngá/ [jíŋgá,jíŋą́] 'crushed grain').

1.2.1.2 Implosive Consonants:

There are two implosive consonants: the bilabial implosive /ɓ/, and the alveolar /ɗ/. The contrast between them and their non-implosive counterparts is most evident in initial position:

(4)

a)	ɗɔ̄ɔ	**to fight**	dɔ̀ɔ̄	**to tie**
	ɗér	**ritual speech**	dér	**pigeon**
	ɗà	**to stand**	dà	**to show**
b)	ɓày	**get up**	bày	**clay bowl**
	ɓèlò	**slave**	bèlò	**hole**
	ɓīi	**sleep**	bìī	**mortar**

In intervocalic position, many speakers pronounce /ɓ/ as [w], while /ɓ/ remains unchanged for others (e.g. /túɓō~túwō/ 'granary cover', /tēɓā~tūwā/ 'chase'). Note that while /b/ at times also alternates with /w/ in this fashion (e.g. /kàw̄~kàbē/ 'type of insect'), in some cases it does not (e.g. /tùbò/ 'lion'). Hence, there are clear contrasts between /ɓ/ and /b/ in intervocalic position (e.g. /túɓō/ 'granary cover' vs. /tùbò/ 'lion').

The implosive /ɗ/ is pronounced /r/ by some speakers, and most especially by younger city or town dwellers. For most Mbay speakers, the distinction between /ɗ/ and /r/ is lost when not at the beginning of a morpheme. A comparison of related languages is needed to reveal the source of non-initial /r/:

(5)

	Mbay	Daba	Gloss
a)	hòr	hòɗè	**fire**
	ndɔ́r̄	ndɔ́ɗɔ̄	**pull**
	ɔ̀r̄	ɔ̀ɗɔ̄	**remove**
b)	ùr	ùr	**dig**
	dàrą̄	dàrā	**sky**
	ngàr̄	ngàrā	**thorn bush**

1.2.1.3 Diphthongs:

In general, the glides /y/ and /w/ form diphthongs quite freely with vowels:

(6)

/y/	iy	wíy	**ululating**
	ey	[rare]	
	ay	Mbáȳ	**Mbay**
		kāy	**split**
	oy	òy	**die**
		dà-bóy	**last**

	ɔy	ɔ̀ȳ	**be heavy**
		kɔ̀ȳ	**owl**
	uy	rùy	**supple**
/w/	iw	tīw	**spit**
	aw	ngàw	**husband**
	ew	ndèw	**tremble**
	ow	tōw	**flute**
	ɔw	ndɔ̄w	**crackle**
	uw	[rare]	

The diphthongs /ay/, /aw/, /oy/, and /ɔy/ occur with far greater frequency than do the others. The diphthong /ɔy/ is pronounced like the 'oy' in English 'boy', and contrasts with /oy/, where the lips are rounded.

As noted above, /w/ is frequently a dialectal alternant to /b/ or /ɓ/. When a schwa follows /b/ or /ɓ/, it is normally dropped in the form containing /w/:

(7)

kàbē ~ kàw̄	**type of bug**
sībē - sēwē	**grey rat**
lábē ~ láw̄	**type of dance**
njéɓā ~ njúwā	**bother**

1.2.1.4 Other Comments:

The phoneme /p/ is relatively rare, occurring almost exclusively in ideophones and in words of foreign origin. The phoneme /h/ is found only at the beginning of a morpheme (e.g. /hòr/ 'fire', /hàl/ 'orphan').

1.2.2 Vowels

The vowel system of Mbay is as follows:

(8)

i u

e ə o

ɔ

a

The vowel /a/ is pronounced similar to Spanish /a/ or the /a/ in English 'father'. Before the nasals (n,m,ng) it is pronounced similar to the 'u' in the English word 'sun' (e.g. /ngāng/ [ngʌ̄ŋg] 'teeth', /nām/ [nʌ̄m] 'friend'). The /e/ is pronounced like the 'e' in Spanish, or like the 'a' in English 'bay', but without the final glide. The /i/ and /u/ are also pronounced like Spanish /i/ and /u/. However, /i/ is normally weakened to /I/ (the 'i' in English 'sit') when followed by 's', 't', 'd', 'k' or 'r', while 'e' is weakened to [ɛ] (the 'e' in English 'set') when followed by 's, t, d, b, l, r, m,' or 'n'. The same weakening process occurs in those rare cases when /i/ or /e/ occur as short vowels in word final position. Before

/r/, both /ə/ and /i/ are very lax: /ə/ almost sounds like schwa, while /i/ is pronounced either as [I] or [U]. When occurring in the first syllable of bisyllabic roots, /i/ and /u/ are neutralized and their pronunciation will be discussed in section 2.2.3 below.

The Mbay /o/ is similar to Spanish /o/ or to the first /o/ in English 'coconut'.[3] It contrasts with /ɔ/, which is pronounced somewhat similar to the 'aw' in English 'law', but shorter and without a final glide. While the distinction between /ɔ/ and /o/ is fairly clear in open syllables, English speakers must take care to distinguish them before /r/ and /y/.

(9)

a)	kɔ̀ȳ	**owl**	kòy	**death**
	gɔ̀y	**type of grass**	gòy	**pestle**
b)	kɔ̄r	**field**	kór	**sesame**
	bɔ̀r̄	**wipe**	bòr̄	**boar**

The distinction between /o/ and /ɔ/ is lost when immediately followed by a nasal (m, n or ng) in the same morpheme.[4] The schwa /ə/ is normally pronounced like the 'e' in English 'tenacious', or like the 'e' in French 'que'. It is pronounced /i/ (like the 'i' in English 'macaroni') when immediately preceding /y/. Before /w/ it is often pronounced [u] or [U].

Foreign words often do not undergo the same changes in pronunciation as do native words. Thus, for example, the /i/ in /àspèríǹ/ 'aspirine', /màsīn/ 'motor', /kìníǹ/ 'quinine', etc., is not weakened to [I].

1.2.2.1 Nasal Vowels:

For all the vowels except for /ɔ/ and /ə/, there is a clear contrast between an oral and a nasal form:[5]

(10)

a)	tàā	**take**	tá̱a̱	**sauce**
	òo	**forge**	ò̱o̱	**burn; eat**
	dèē	**person**	dè̱é̱	**woman**
b)	kérā	**one**	kérā̱	**accuse**
	bíyā	**porridge**	bīyā̱	**goat**
c)	kàr̄	**clearness**	kà̱r̄	**toad/frog**
	bàr	**plank**	bà̱r̄	**pat. aunt**

The contrast between nasalized vowels and oral vowels is most common in open monosyllabic morphemes. Outside of their occurrence in morphemes of this type, they are only found adjacent to /n/, /m/, /ng/, /y/ and /r/. (This is the primary reason CCL has proposed two distinct phonemes nasalized /y̱/ and /ṟ/.)

There is no contrast beween nasalized /ǫ/ and nasalized /ɔ̨/: although they are pronounced /ɔ̨/ in an open syllable, I follow CCL in using the symbol /ǫ/ to represent this sound.

The contrast between oral and nasalized vowels is lost when:
1) a vowel precedes a nasal consonant. In this case, it is always pronounced nasalized. For example, /yā̆m/ 'be thin' and /nā̆m/ 'friend' are pronounced [yą̄m] and [ną̄m]. My transcription does not include the nasalization. Since there exists no distinction between nasalized /ǫ/ and /ɔ̨/, the distinction between /o/ and /ɔ/ is also lost before nasal consonants as well, and either pronunciation is usually possible. However, it is my impression that certain words favor one vowel over the other: for example, /ngō̆n/ 'child' seems to only occur only with /o/, /kɔ́n/ 'nose' only with /ɔ/, while /ndō̆m ~ ndɔ̄̆m/ 'to dance'.[6]
2) a vowel is in final position in a morpheme and follows a nasal consonant. In this case, it is also nasalized:

(11)

mį́į	**five**	kùmą̄	**sorcerer**
ną̄ą	**moon**	kèną̄	**between**
nę̄ę́	**wife**	némą́	**okra**

Note that the rare nasalized vowel /ę/ has an alternant pronunciation of [ǣ̨] in a small number of words (e.g. /ànę̄/ [à·nǣ̨ǣ̨] 'and so, then').
3) when in the first syllable of a bisyllabic word and following a nasal and preceding an obstruent. In this case, it is not pronounced like a nasalized vowel:

(12)

nàsə̄	**antelope**	màsə̄	**be acidic**
nàtə̀	**forehead**	mātə̄	**type of tree**

4) when in the second syllable of a bisyllabic root and following a prenasalized stop. In this case, the vowel is normally not nasalized (e.g. /síndá/ 'horse', /sùmbā̆/ 'aim'), unless the prenasalized stop is /ng/, in which case nasalization can occur with some speakers (e.g. /jíngą́ ~ jíngá ~ jíŋą́/ 'crushed grain'). However, the suffix /-a/ is normally nasalized when it is attached to a root ending in a prenasalized stop followed by schwa (e.g. /kàndə̄-á/ [kàndą́] 'its fruit').

When a vowel is preceded by a nasal and followed by a sonorant in a monosyllabic word, a contrast between oral and nasalized vowels does occur:

(13)

màý	**upside down**	mą̀ȳ	**carefully**
		ną̀y	**be lost**

nàr	money	mà̰r	insist
màr	crocodile	mā̰r	invite

In this work, in most cascs I have normally marked nasalized vowels where they occur, since nasalized vowels occur independently of nasals. However, in bisyllabic words, only one vowel is normally marked as nasal (e.g. /bīyā̰/ [bḭ̄yā̰] 'goat', even though nasalization spreads to all vowels within a morpheme. However, in bisyllabic words ending in a vowel preceded by a nasal, I have also accepted the CCL spelling as an alternant (e.g. /kènā~kènā̰/ 'between').

Note that when suffixes consisting of nasals are added to a root ending in a vowel, the vowel is normally pronounced slightly nasalized. In spite of this fact, neutralization does not occur in this environment: thus, /ìī-ñ/ 'dry up (pl.)' is phonologically and phonetically distinct from /ḭ̀ḭ̄-ñ/ 'they are (pl.)'.

1.2.2.2 Schwa:

The schwa (/ə/) is an exceptional vowel in a number of ways. 1) as noted above, there is no contrast between oral and nasalized schwa; 2) there is no length contrast for schwa: it is always short; 3) when another vowel follows schwa, schwa normally becomes identical to this vowel:

(14)

kə́ à --->	káà	which will
tə́ àí --->	táàí	not in
kāgə̄~á --->	kāgāá	its wood

The distribution of schwa is also quite limited. In native monosyllabic roots, schwa is found exclusively in grammatical formatives: prefixes or suffixes indicating the person of the subject or object (e.g. /kə̀/ 'we', /də́/ 'them'), complementizers (e.g. /kə́/ 'that'), etc. In multi-syllabic words, schwa can be shown to be the result of the neutralization of other vowels. This will be discussed in detail in section 2.

1.2.2.3 Vowel Length

Mbay, unlike most Sara languages, has a clear contrast based on the length of vowels:

(15)

tò	lie	tòo	pirogue
tō	carry	tōo	hurt, pain
mbā	propose	mbā̰a	guest
ndī	voice	ndīi	rain
ɗá	where?	ɗáa	solidly (tie)

Contrasts of this type, which are fairly common, are limited (with very few exceptions) to morphemes consisting of one syllable ending in a vowel.

In fact, the occurrence of long and short vowels in words is somewhat predictable. The distribution of short vowels can be described as follows:

a) both vowels in bisyllabic morphemes are normally short.

b) the vowel in a morpheme ending in a sonorant consonant is always short. There are very few exceptions (e.g. /kàar/ 'pond', /kāar/ shell), and it is my impression that even these are often pronounced short in quick speech.

c) the vowels found in grammatical formatives are normally short. These include all vowels in prefixes and suffixes, vowels in complementizers, vowels in the names of more commonly used parts of body (e.g. /tà/ 'mouth', /jī/ 'hand', /dɔ̀/ 'head', /rō/ 'body', /ndī/ 'voice'), especially those which also function as prepositions and in word-formation processes.

d) vowels in some of the interrogative pronouns are short (e.g. /ɗí/ 'what', /ɗá/ 'where' , but /ná̰ā̰/ 'who?').

e) vowels found in auxiliary verbs are short (e.g. the progressive markers /ndì/,/ ɗà/,/tò/, the future marker /à/, and /ì/ 'to be').

f) a large class of verbs which are derived historically from bisyllabic morphemes beginning with a vowel have lost the initial vowel and now contain a short vowel in an open monosyllabic morpheme:

(16)

	MBAY	SAR	DABA	
a)	là	èlà	ùlà	to send
	ndà	èndà	ùndà	to hit
	ndā	èndā	ūndā	to set, put
b)	jō	èjō	ìjō	to weave
	tō	ètō	ìtō	to carry

Long vowels are found almost exclusively in morphemes of a single syllable ending in a vowel. Further, with the exception of grammatical formatives and truncated bisyllabic roots as noted above, all morphemes of this shape contain long vowels. The same exceptions exist (e.g. /kāar/ 'type of shell', /kàar/ 'pond', /kɔ̄ɔr/ 'ball of food'), but they are very rare.

The transcription of vowel length found in Caprile (1968) differs in an important way from that used here. Morphemes of the shape CVC which have two different tones are marked long by Caprile (e.g. /bòr̄/ 'boar' is written /bòōr/), while those with two identical tones (e.g. /kày/ 'hartebeest') are left unchanged. There are a number of problems with this approach. First, the transcription suggests a difference in the length of the vowel in such cases where none exists. Further, in those rare cases where the vowel is long, it contrasts in length with bi-tonal morphemes (e.g. /kàar/ 'type of shell' vs. /kàr̄/ 'clearness'), but his transcription will not allow for this distinction. Further, there is substantial evidence, both historical and phonological (see below), that forms such as /kày/ and /bòr̄/ are actually derived from /kàyà/ and /bòrō/, and that the final vowel has been dropped and the tone left on the sonorant. For these reasons, in this work I follow instead the system of transcription found in Magriña and Fournier (1974).

In addition to the length contrast noted, there exists at the phonetic level a third vowel length, extra long, which occurs in a fairly large class of ideophones. These are formed by the addition of a vowel suffix, identical in quality to the stem vowel. As a result, contrasts such as /dòo/ 'wound' vs. /dóoo/ 'perfectly (white)' occur. Occasionally a three-way contrast occurs (e.g. /sō/ 'to fall' /sōó/ 'here', /sóoo/ 'completely (pierced)'). However, extra long vowels are found only in ideophones; the phonetic length is generally longer than the transcription suggests, and varies substantially, depending on the emphasis the speaker wishes to give to it.

1.3 Loan Words

Mbay contains a large number of words borrowed from other languages, especially French, Arabic, and Sango. Words from French include most terms related to government, administration, education, and modern tools and technology. Arabic words include many tools, weapons, garments and religious terms introduced by the Arabs. Almost all the words related to horse-riding are of Arabic origin. Sango words include a large of number of tuber species and foods made from them, as well as a number of bird and animal species of Central African origin.

Foreign sounds which do not exist in Mbay are replaced in a fairly regular fashion. Words containing French /ch/ [š] appear in Mbay with /s/:

(17)

FRENCH	GLOSS	MBAY	GLOSS
charpentier	carpenter	sàrbàndī	board
changer	change	sānjē	change
chapeau	hat	sàpóò	hat

Foreign words containing /f/ appear in Mbay with /p/, while French words with /v/ normally appear with /b/ or /w/:

(18)

FRENCH	GLOSS	MBAY	GLOSS
chef	chlef	sépè	chief
fleur	flower	pèlér̀	flower tree
serviette	napkin	sèrbétè	towel
civil	civilian	sùwél̀	civilian

Educated speakers of Mbay will tend to retain the French pronunciation of these sounds.

Many French words which are normally used with a definite article (`l´`, `le`, or `la`) appear in Mbay with an initial /`l`/:

(19)

lòkól̀	(l´école)	**school**
làmpǫ́ǫ̀	(l´impôt)	**taxes**
lòpìtál̀	(l´hôpital)	**hospital**
làmsǫ́ǫ̀	(la mission)	**mission**
làkér	(la clé)	**key**

Arabic words containing an /`m`/ between two vowels, where stress immediately precedes the /`m`/, appear in Mbay with /`mb`/:

(20)

ARABIC	MBAY	
tumur	túmbùr	**date**
gumar	gúmbàr	**card game**
jamal	jámbàl	**camel**

When a foreign word contains a voiced sibilant such as /`z`/, or a voiceless /`s`/ which is affected by neighboring sounds or which is pronounced differently in the source language (e.g. Arabic /`ṣ`/), it often appears in Mbay as /`j`/:

(21)

MBAY	GLOSS	FROM	
sànjèmą̄	**census**	recensement	**(French)**
pàránjè	**French**	français	**(French)**
sèrjōo	**bike rim**	cerceau	**(French)**
jéy	**true**	sey (ṣaḥi)	**(Arabic)**
sèwàjī	**used clothing**	choisir	**(French)**
séjè	**chair**	chaise	**(French)**

1.4 Tonal System

There are three level tones in Mbay: low (`à`), mid (`ā`), and high (`á`). Tone-bearing elements are vowels (including both elements of long vowels), sonorant consonants which occur at the end of a syllable, and syllabic nasals:

(22)

H	ké	**one which**	ń	**the one which**
M	kē	**towards**	m̄-	**I**
L	kè	**with**	ǹ-	**he (indirect form)**
HH	súu	**spider**	kɔ́n	**nose**
	mérą́	**taste**	kádé	**stalk**
HM	sáā	**look for**	kán̄	**to assemble**
	téndā	**to beat(freq.)**	látē	**to rub on**
HL	sóò	**bucket**	túm̀	**garlic**
			látè	**badly (scratch)**
MH	tą̄ą́	**first**	ndāŕ	**skin**

	ndījá	**scars**	mōgó	**type of fish**
MM	kōo	**to shout**	ngōn	**child**
	bīyā̰	**goat**	mātē	**type of tree**
ML	ngār̀	**chief, king**		
LH	kàá	**even**	gàḿ	**snare**
	kùndé	**millet drink**	gògé	**outside**
LM	tàā	**to take**	màn̄	**water**
	ndìsā	**broom**	bàtē	**sheep**
LL	tòo	**pirogue**	hàl	**orphan**
	bìyà̰	**type of bee**	lìtè	**falling apart**

Of the nine sequences possible on bi-tonal morphemes, seven are common, while two, high-low and mid-low, are restricted. The sequence high-low occurs very rarely in native vocabulary except with descriptives (ideophones). It is common in loan words, especially those from French and Arabic.

(23)

French:			
àsétè	**plate**	sàmàpúù	**I don't care**
séjè	**chair**	bíkè	**pen**
Arabic:			
ítèr	**perfume**	kàlásè	**already**
tábàkè	**straw plate**	ásà	**now**

The sequence mid-low is extremely rare, found only in very few morphemes (e.g. /ngār̀/ 'chief').

In a handful of cases, complex tones (two tones together) are found on a single tone-bearing element (e.g. /ň/ 'with it/them', /nà̰-tâ/ 'otherwise').

The system of transcription for marking tones used in this work is taken directly from that of CCL. Each tone-bearing element is marked for tone except when it carries the same tone as the segment which immediately precedes it (e.g. /tàa/ 'words' vs. /tàā/ 'to take' and /àl/ 'to swim' vs. /àl̄/ 'to climb').

The two restrictions on tonal sequences noted above (i.e. rarity of mid-low and high-low sequences in native vocabulary) operate only at the morphemic level, and have no effect on morphologically complex words (e.g. /lā-ǹ/ 'throw to him').